NON-CANONICAL MICRORNAS: CHARACTERIZATION OF THE HUMAN MIRTRON BIOGENESIS PATHWAY

Ph.D. Thesis

Anita Schamberger

EÖTVÖS LORÁND UNIVERSITY, FACULTY OF SCIENCE

DOCTORAL SCHOOL OF BIOLOGY

CLASSICAL AND MOLECULAR GENETICS DOCTORAL PROGRAM

Head of Doctoral School:	dr. Anna Erdei
Head of Doctoral Program:	dr. Tibor Vellai
Supervisor:	dr. Balázs Sarkadi
Consultant:	dr. Tamás Orbán

Institute of Molecular Pharmacology, Research Centre for Natural Sciences, Hungarian Academy of Sciences

Budapest

2013

INTRODUCTION

Since the discovery of the first small regulatory RNA molecules (Lee et al., 1993; Wightman et al., 1993), microRNAs (miRNAs) have become one of the most studied fields in molecular biology. These small non-coding RNA molecules form a gene regulatory network comparable to that of transcription factors by targeting mRNA molecules via sequence complementarity (Filipowicz et al., 2008; Carthew & Sontheimer, 2009; Ghildiyal & Zamore, 2009; Slezak-Prochazka et al., 2010). Most of these 20–24 nucleotide (nt) long molecules are formed via a canonical pathway, beginning with the transcription of a long primary precursor (pri-miRNA) with hairpin structure, followed by the cleavage of the nuclear RNase III-like enzyme Drosha, mediated by its partner protein, DGCR8. The so formed pre-miRNA is then recognized and transported by the Exportin-5 shuttle system to the cytoplasm where Dicer, another RNase III-type enzyme, produces the double-stranded miRNA:miRNA* duplex. Further processing involves the incorporation of this species into an Argonaute protein containing RNA-induced silencing complex (RISC) where one strand is removed and the mature RISC is formed. The regulatory effect is then manifested by the destabilization and/or translational inhibition of the target mRNA molecule via partial base pairing of the miRNA and the 3'-untranslated region (3'-UTR) of the mRNA (Huntzinger & Izaurralde, 2011).

Owing to the development of molecular biology techniques and bioinformatics tools, there are emerging data on various alternative miRNA maturation routes beside the canonical miRNA biogenesis pathway. The alternative biogenesis pathways could bypass certain steps of the canonical process, typically one of the two cleavage steps. Therefore, there are Drosha-independent and Dicer-independent alternative maturation pathways (Miyoshi et al., 2010; Yang&Lai, 2011).

The most prominent Drosha-independent miRNA biogenesis is the mirtron pathway which was first described in *Drosophila melanogaster* and *Caenorhabditis elegans* (Okamura et al., 2007; Ruby et al., 2007). MiRNAs of mirtron origin are localized in short introns (50-100nt) where the whole intron is essentially equivalent to the pre-miRNA form. This allows that the first step of the mirtron processing is different from the canonical one: the pre-miRNA is cleaved out from the primary transcript by the splicing machinery instead of the Drosha/DGCR8 complex.

Beside invertebrates, bioinformatics studies and small RNA data analysis suggested the existence of the mirtron biogenesis pathway also in vertebrates (Berezikov et al., 2007; Babiarz et al., 2008; Glazov et al., 2009; Chong et al., 2010; Babiarz et al., 2011). Moreover, among other mammalian species, mirtrons were predicted to be present also in humans, but experimental validations or detailed investigations in these species were still remained.

AIMS

One of the most widespread non-canonical miRNA maturation routes is the mirtron biogenesis pathway. It was discovered in invertebrates, but was also predicted in vertebrates, e.g. in mammals. Since at the beginning of my work there was no experimental evidence for the existence of the mirtron biogenesis pathway in these species, we decided to investigate the following:

- Processing of functional mature miRNAs from three predicted human and one *Drosophila* (control) mirtron from their original genomic coding context in human cells.

- Investigation of the effect of the harboring exon sequences on mature miRNA processing for the chosen mirtrons, analysis of their expression and function from a heterologous coding context.

- Development of stable mirtron-expressing cell lines by using the *Sleeping Beauty* transposon based gene delivery system.

- To provide evidence for the mirtron biogenesis pathway in mammals: investigation of splicing dependency and Drosha/DGCR8 independency of miRNA processing from predicted mirtron origin.

- Direct detection of mature miRNAs of mirtron origin.

METHODS

- We have made more than fifty plasmid constructs for our experiments (mirtron/ miRNA expressing and control vectors, luciferase sensor/ mutant sensor plasmids, etc.). They were made by standard molecular cloning methodologies. Subsequently, the transfection efficiencies of these vectors were followed by EGFP fluorescence.

- The applied mirtron, control and exon coding sequences were selected by various bioinformatics tools, available free of charge on the internet.

- Splicing of the predicted mirtron sequences were examined 48h after transfection by the isolation, reverse transcription and PCR (RT-PCR) of total RNA samples. Accurate splicing was verified by sequencing the PCR products.

- The function of mirtron derived mature miRNAs was detected by luciferase assays, applying sensor and mutant sensor constructs, as well as ’non-cognate’ miRNA controls.

- We established stable mirtron- and canonical miRNA-expressing cell lines by the *Sleeping Beauty* transposon based gene delivery system. We used EGFP fluorescence for sorting the cells.

- We have confirmed the existence of mirtron biogenesis pathway in two different ways. First, we proved splicing dependency by making and investigate (RT-PCR, luciferase assay) 5’ splicing mutant mirtron expressing vectors. Second, we demonstrated Drosha/DGCR8 independence by luciferase assays in DGCR8 deficient mouse embryonic fibroblast cells.

- For the direct detection of human mirtron derived mature miRNAs, we applied Northern blot and stem-loop quantitative RT-PCR techniques.

RESULTS

- We have found that the investigated predicted human mirtrons were able to splice out from their original genomic environment. However, the splicing efficiencies were different and generally low for each mirtron construct. In spite of the good splicing predictions, the *Drosophila* mirtron control was not able to splice out from its harboring genomic exons in mammalian cells.
- Out of the three investigated predicted mirtrons, only two (mir-1226 and mir-877) could generate functional mature miRNAs, when expressed from their original genomic context.
- When expressing predicted mirtrons from a heterologous coding context, in the case of mir-1226 and mir-877, we got similar results to that of mirtrons expressed from endogenous coding context.
- We successfully developed stable mirtron-expressing cell lines by the *Sleeping Beauty* transposon based gene delivery system. We did not find any conspicuous difference among the developed, mirtron overexpressing and the parental cell lines during normal cell culture conditions. These data indicate that expression of the investigated mirtrons does not interfere with the expression of essential genes.
- We have found that mir-877 can liberate functional mature miRNAs from both of its 5'- and 3'-arms simultaneously in various cell types examined. The miR-877-3p repressed the luciferase sensor significantly, and it was comparable to that of miR-877-5p.
- The 5' splice site mutation of mir-1226 and mir-877 abolished the generation of functional miRNAs from these sequences. It has proved the splicing dependency of mature miRNA processing from the mir-1226 and mir-877 mirtrons.
- In contrast to the canonical miRNA control, functional mature miRNAs were generated from mir-877 in DGCR8 deficient mouse embryonic fibroblast cells. It provided evidence that there is Drosha/DGCR8 complex independent mirtron biogenesis in mammals.
- The pre- and mature miRNA forms of mir-1226 were detectable by Northern blot analysis, while mir-877 derived miRNA forms were not.
- For the reliable detection of mir-877 derived mature miRNAs, we optimized the stem-loop quantitative RT-PCR technique. During these experiments:

- we determined the optimal amount of the total RNA for the reverse transcription reaction for various endogenous controls;
- we have shown that the target miRNA of interest can be reverse transcribed together with the appropriate endogenous control in one reaction;
- we demonstrated that carry over DNA contamination could severely mislead the detection by this technique;
- we concluded that the effect of the DNA on the measurements is sequence specific, since only DNA containing the corresponding pre-miRNA coding sequence gave false positive signals;
- our data revealed that the false positive signal is mostly realized during the reverse transcription reaction.

- After a proper optimization and establishment of a refined protocol of the stem-loop quantitative RT-PCR, we reliably detected all of the mirtron derived functional mature miRNAs by this technique. miR-1226 was detectable at a level of one order of magnitude higher than miR-877-5p or mir-877-3p.

CONCLUSIONS

We provided evidence that the splicing dependent and Drosha/DGCR8 complex independent mirtron biogenesis pathway exist not only in invertebrates but also in vertebrates. There are human mirtron derived functional miRNAs, and their processing is independent of the flanking exons. However, experimental validation revealed that there are predicted mirtron sequences which are not *bona fide* mirtrons. Furthermore, we showed that mir-877 can liberate functional mature miRNAs from both of its 5'- and 3'-arms simultaneously in various cell types examined. Therefore, we would like to emphasize that it seems to be important to examine both arms of a particular miRNA when exploring the relevant physiological functions. Moreover, our data suggest that neither good splicing efficiency nor high overall miRNA expression level is a strict prerequisite for efficient mirtron functionality.

Optimization of the stem-loop quantitative RT-PCR showed a surprising result that DNA could serve as a template during mature miRNA measurements, mostly during the reverse transcription reaction. Our data suggest that the corresponding pre-miRNA coding sequence is detected by the stem-loop primer. Our data that the mir-1226 derived mature

miRNA was detected one order of magnitude higher than that of mir-877 by this technique could explain why mir-877 derived mature miRNA forms could not be detected by Northern blot analysis.

REFERENCES

Babiarz JE, Hsu R, Melton C, Thomas M, Ullian EM, Blelloch R. 2011. A role for noncanonical microRNAs in the mammalian brain revealed by phenotypic differences in Dgcr8 versus Dicer1 knockouts and small RNA sequencing. *RNA 17*:1489-1501.

Babiarz JE, Ruby JG, Wang Y, Bartel DP, Blelloch R. 2008. Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs. *Genes Dev 22*:2773-2785.

Berezikov E, Chung WJ, Willis J, Cuppen E, Lai EC. 2007. Mammalian mirtron genes. *Mol Cell 28*:328-336.

Carthew RW, Sontheimer EJ. 2009. Origins and Mechanisms of miRNAs and siRNAs. *Cell 136*:642-655.

Chong MM, Zhang G, Cheloufi S, Neubert TA, Hannon GJ, Littman DR. 2010. Canonical and alternate functions of the microRNA biogenesis machinery. *Genes Dev 24*:1951-1960.

Filipowicz W, Bhattacharyya SN, Sonenberg N. 2008. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nat Rev Genet 9*:102-114.

Ghildiyal M, Zamore PD. 2009. Small silencing RNAs: an expanding universe. *Nat Rev Genet 10*:94-108.

Glazov EA, Kongsuwan K, Assavalapsakul W, Horwood PF, Mitter N, Mahony TJ. 2009. Repertoire of bovine miRNA and miRNA-like small regulatory RNAs expressed upon viral infection. *PLoS One 4*:e6349.

Huntzinger E, Izaurralde E. 2011. Gene silencing by microRNAs: contributions of translational repression and mRNA decay. *Nat Rev Genet 12*:99-110.

Lee RC, Feinbaum RL, Ambros V. 1993. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell 75*:843-854.

Miyoshi K, Miyoshi T, Siomi H. 2010. Many ways to generate microRNA-like small RNAs: non-canonical pathways for microRNA production. *Mol Genet Genomics 284*:95-103.

Okamura K, Hagen JW, Duan H, Tyler DM, Lai EC. 2007. The mirtron pathway generates microRNA-class regulatory RNAs in Drosophila. *Cell 130*:89-100.

Ruby JG, Jan CH, Bartel DP. 2007. Intronic microRNA precursors that bypass Drosha processing. *Nature 448*:83-86.

Slezak-Prochazka I, Durmus S, Kroesen BJ, van den Berg A. 2010. MicroRNAs, macrocontrol: regulation of miRNA processing. *RNA 16*:1087-1095.

Wightman B, Ha I, Ruvkun G. 1993. Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. *Cell 75*:855-862.

Yang JS, Lai EC. 2011. Alternative miRNA Biogenesis Pathways and the Interpretation of Core miRNA Pathway Mutants. *Mol Cell 43*:892-903.

PUBLICATIONS RELATED TO THE THESIS

Accepted publications in peer-reviewed journals

Schamberger A and Orbán TI: Experimental validation of predicted mammalian microRNAs of mirtron origin. *Methods Mol Biol*: *Structural and Functional RNA Mapping,* Springer Verlag book (edited by Lucrecia Alvarez)
(book chapter; accepted)

Schamberger A, Sarkadi B, Orban TI: Human mirtrons can express functional microRNAs simultaneously from both arms in a flanking exon-independent manner. *RNA Biology*. 2012, 9(9):1177-85.
IF : 4,93

Orbán TI, Apáti A, Németh A, Varga N, Krizsik V, Schamberger A, Szebényi K, Erdei Z, Várady G, Karászi E, Homolya L, Német K, Gócza E, Miskey C, Mátés L, Ivics Z, Izsvák Z, Sarkadi B: Applying a "double-feature" promoter to identify cardiomyocytes differentiated from human embryonic stem cells following transposon-based gene delivery. *Stem Cells***.** 2009, 27(5):1077-87.
IF: 7,75

Under publication in peer-review journals

Schamberger A and Orbán TI: 3' isomiR species and DNA contamination influence reliable quantification of microRNAs by stem-loop quantitative PCR (under publication)

Oral conference presentations

Schamberger A, Sarkadi B, Orbán TI: Non-canonical human microRNAs: characterization of the mirtron pathway in higher eukaryotes. Hungarian Molecular Life Sciences; 5-7 April, 2013, Siófok, Hungary. Abstract book: O-080.

Poster conference presentations with first authorship

Anita Schamberger and Tamás I. Orbán: Investigation of the miRNA stem-loop quantitative PCR: DNA contamination and 3' isomiR species might influence reliable quantification. EMBO|EMBL Symposium, The Non-Coding Genome, 9-12 October, 2013, Heidelberg, Germany. Abstract book: 293.

Anita Schamberger, Balázs Sarkadi, Tamás I. Orbán: Experimental validation of human mirtrons: functional microRNAs can be produced simultaneously from both arms in a flanking exon-independent manner. EMBO|EMBL Symposium, The Complex Life of mRNA; 7-10 October, 2012, Heidelberg, Germany. Abstract book: 248.

Schamberger A., Kasza I., Homolya L., Sarkadi B., Orbán T.: Sleeping Beauty transzpozon alapú génbevitel: GFP-címkézett ABCG2 multidrogtranszporter-variánsok karakterizálása stabil sejtvonalakban. Biokémia, 2009, XXXIII. (3): 38.

A. Schamberger, A. Apati, N. Varga, V. Krizsik, A. Nemeth, Z. Izsvak, C. Miskey, Z. Ivics, K. Nemet, B. Sarkadi, and T. I. Orban: Stable transgene integration into human embryonic stem cells using the Sleeping Beauty transposon system. 6th Annual International Conference on Transposition and Animal Biotechnology, June 19-21, 2008, Berlin, Germany. Abstract book: 83.

OTHER PUBLICATIONS IN PEER-REVIEWED JOURNALS

Kolacsek O, Krízsik V, Schamberger A, Erdei Z, Apáti Á, Várady G, Mátés L, Izsvák Z, Ivics Z, Sarkadi B, Orbán TI: Reliable transgene-independent method for determining Sleeping Beauty transposon copy numbers. *Mobile DNA*. 2011, 2(1):5.
IF (predicted): 2,64

Sarkadi B, Orbán TI, Szakács G, Várady G, Schamberger A, Erdei Z, Szebényi K, Homolya L, Apáti A: Evaluation of ABCG2 expression in human embryonic stem cells: crossing the same river twice? *Stem Cells*. 2010, 28(1):174-6.
IF: 7,87

Apáti A, Orbán TI, Varga N, Németh A, Schamberger A, Krizsik V, Erdélyi-Belle B, Homolya L, Várady G, Padányi R, Karászi É, Kemna EWM, Német K, Sarkadi B: High level functional expression of the ABCG2 multidrug transporter in undifferentiated human embryonic stem cells. *Biochimica et Biophysica Acta - Biomembranes*. 2008, 1778(12): 2700-9.
IF: 4,18

Szendrői A, Schamberger A, Bakonyi T, Hornyák A, Rusvai M: Molecular comparison of the fibre proteins of bovine adenovirus subtype A and subtype B. *Virus Res*. 2003, 94(1): 59-63.
IF: 1,72